AVIAN MUSING FEATURE SPACE ANALYSIS

A Thesis
Presented to
The Academic Faculty

by

Guillermo J. Colón

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
August 2012

AVIAN MUSING FEATURE SPACE ANALYSIS

Approved by:

Professor David V Anderson,
Committee Chair
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor David V Anderson, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Patricio Vela
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Justin Romberg
School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: 17 May 2012

To my wife,

Dorthea M. Colón,

and my parents,

Olvia E. Villafañe and Guillermo Colón,

for their continual support.

ACKNOWLEDGEMENTS

I am very grateful for the great and constant support of my advisor Dr. David V. Anderson, and the collaboration with some of my colleages working on different aspects of this project, such as Nathan V. Parrish and Ryan Curtin.

I would also like to acknowledge the support of the State of Georgia that has been sponsored through the Agricultural Technology Research Program at Georgia Tech in cooperation with the Georgia Poultry Federation and its member companies the poultry-related work described herein. Also, thanks to the many contributions of my colleagues in the Food Processing Technology Division of the Aerospace, Transportation and Advanced Systems Lab in the Georgia Tech Research Institute.

I would like to especially acknowledge Wayne Daley, Sim Harbert, Doug Britton, Matthew Giannelli, and Erin Hanson. Also for the support of the administrative assistants such as Lucy D. Johnson, and other people in supportive roles. I am very grateful for all of the people that helped us perform the experiments over at the University of Georgia at Athens.

At UGA we have been working with the Department of Poultry Science with, advice/guidance provided by Drs. Bruce Webster and Casey Ritz. In addition I would like to thank the management and staff at the Poultry Research Farm for their assistance and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The purpose of this study was to analyze the possibility of utilizing known signal processing and machine learning algorithms to correlate environmental data to chicken vocalizations. The specific musing to be analyzed consist of not just one chicken's vocalizations but of a whole collective, it therefore becomes a chatter problem. There have been similar attempts to create such a correlation in the past but with singled out birds instead of a multitude. This study was performed on broiler chickens (birds used in meat production).

One of the reasons why this correlation is useful is for the purpose of an automated control system. Utilizing the chickens own vocalization to determine the temperature, the humidity, the levels of ammonia among other environmental factors, reduces, and might even remove, the need for sophisticated sensors.

Another factor that this study wanted to correlate was stress in the chickens to their vocalization. This has great implications in animal welfare, to guarantee that the animals are being properly take care off. Also, it has been shown that the meat of non-stressed chickens is of much better quality than the opposite.

Following this introduction, this stress will be described in detail, along with its consequences. It will cover animal welfare issues. Further, the experiment design will be explained, and the issues of noise will be covered. Then, feature space is explored and discussed. Finally the results are analyzed using different statistical tools.

CHAPTER II

PROBLEM BACKGROUND

2.1 Animal Welfare

Through the years there has been an increasing awareness of the need for animal welfare. This not only from pro-animal rights groups, but from the poultry industry itself. Research has shown the need for animal welfare; it has shown that "happy" or unstressed birds not only produce more meat, but better quality meat [11]. This improves business, which makes the industry "happy" or unstressed. This is why the poultry industry strives to ensure the good treatment of the animals they produce.

What is stress? "Stress can be defined as the set of responses to external demands which calls upon the flocks to adapt to a new or abnormal situation [4, 7, 11]. This process of adaptation causes the release of hormones and requires the redistribution of body reserves including energy and protein at the cost of decreased growth, reproduction, and health [3, 11]. After extended or repeated periods of stress, birds become fatigued and weak; they often succumb to starvation and infectious diseases [4, 6, 11]."

2.2 Causes of Stress

There are several stress factors that affect the birds and therefore produce less meat. These are classified into several categories of stress: climatic, environmental, nutritional, physiological, physical, social and psychological. In this research we focused on using climatic and environmental stress factors, heat and ammonia [11].

2.2.1 Heat

Chickens have a normal body temperature of 41.1°C and are the "most" comfortable in 10°C to 20°C temperatures. According to Rural Chemical Industries research, stress induced by heat is the major cause of loss of profits and production in the hotter areas of the world. These hotter areas are places with temperatures above 25°C and a relative humidity above 40%. Birds under this type of stress, drink more water and eat less (their appetite goes down by 1.5% for each centrigrade above 20°C.) The more water the birds drink the more they lose electrolytes. The core temperature of the birds is higher as they are stressed or sick. They drink more, they have a higher respiration rate, lower pulse rate, and the amount of wet droppings increases (this in turn hastens dehydration). It is a very destructive pattern that ruins production and culminates in death [1].

2.2.2 Ammonia

Atmospheric ammonia is produced in broiler pens due to the excretions of the birds, they release uric acid ($C_5H_4N_4O_3$) as a way to remove ammonia from their bodies (produced from their digestive process), this uric acid breaks down once outside the body into ammonia again. The concentration of birds in the pen makes the ammonia levels produced very high. Miles showed that "broilers exposed to concentrations greater than 25 ppm of atmospheric ammonia experienced a reduction in body weight, and generally had greater mortality." The manner in which this contaminant can be reduced is through increased ventilation but this "compromises energy efficiency" [9].

2.3 Current Methods to Control Stress

The industry has no direct measurement of stress other than an intrusive blood test. The bird's stress is typically monitored and attempted to be controlled by means of the environment. The temperature is controlled to the "most" comfortable setting by means of heating and cooling elements, which are monitored by single point thermal sensors. The humidity is controlled by means of ventilation, and is typically monitored by single point humidity sensors. The ammonia levels are also controlled by means of ventilation, most often timed or turned on manually by a worker (who determined the necessity for it by smell, or sometimes checked by very sensitive equipment).

All of these methods focus on evaluating and controlling the environment and are not directly detecting and controlling stress. What is novel about this approach is that the end-goal is to have a system that attempts to monitor the stress in the birds directly and not the environmental variables.

CHAPTER III

EXPERIMENT DESIGN

All of the bird experiments were performed at the University of Georgia in Athens, and in compliance with federal requirements (USDA, OLAW) and IACUC practices. There were two main setups for the experiments, a full grow-out room and a smaller chamber. The full grow-out room allowed acquisition of data over the life of a large number of birds. The chamber setup allowed for a more controlled environment, with fewer birds and short-term data acquisition. The chamber and the computer system can be observed in Figure 1.

3.1 The Hardware

The data acquisition system was custom-made for the experiments. The system consisted of an Intel Q8400 processor (quad-core 2.66GHz with 4M cache) with 4GB of RAM and four Western Digital Caviar Black WD1001FALS 1TB 7200 RPM drives set up in a soft RAID5 array, running Debian Linux. The audio itself was captured by four Shure KSM141/SL Multi-Pattern Condenser Studio Microphones connected to the computer via Shure X2u XLR to USB Microphone Signal Adapters. The environmental data was acquired through SparkFun's USB Weather Board (SEN-09800) which has a SCP1000 temperature and barometric pressure sensor, the TEMT6000 ambient light sensor, and the SHT15 temperature and humidity sensor. In addition to this USB Weather Board, more environmental data (for validation) was obtained by two Vaisala HUMITTER® Temperature and Humidity sensors connected through a LabJack U12i interface.

Figure 1: Data Acquisition System and Grow-out Chamber

3.2 The Software

Two software packages were created for the data acquisition in this project, one to record long-term and one to record short-term experiments. The main difference lies in usability, how the data is stored and how the system is configured.

In the first package the audio is recorded using SoX with cron scheduling it to run every minute. This is done by means of simple shell scripts. The main problem with this setup is that a few frames get dropped ever so often. The number of frames skipped was minimal, so for the time being, this problem was ignored. The system provided a simple ncurses (text) terminal interface, for logging visits into the coop. There are custom programs that capture the environmental data into CSV files and it is scheduled through cron also.

Figure 2: The birds in experiment chamber

The second package provided a more elaborate interface, including an data monitor, as can be observed in Figure 3. The system consisted of one startup script which initiates all data acquisition. The audio is acquired by an in-house program that captures the audio directly through the ALSA interface. That program communicates with the monitoring program by means of shared memory through IPC. The environmental data is captured by the same programs as in the first package, but they can communicate with the monitoring program in a similar manner as the audio capture program (by means of IPC shared memory).

3.3 Experiments

There were three notable experiment sets:

Full Room Grow-out - October through December 2009 Stressed was induced

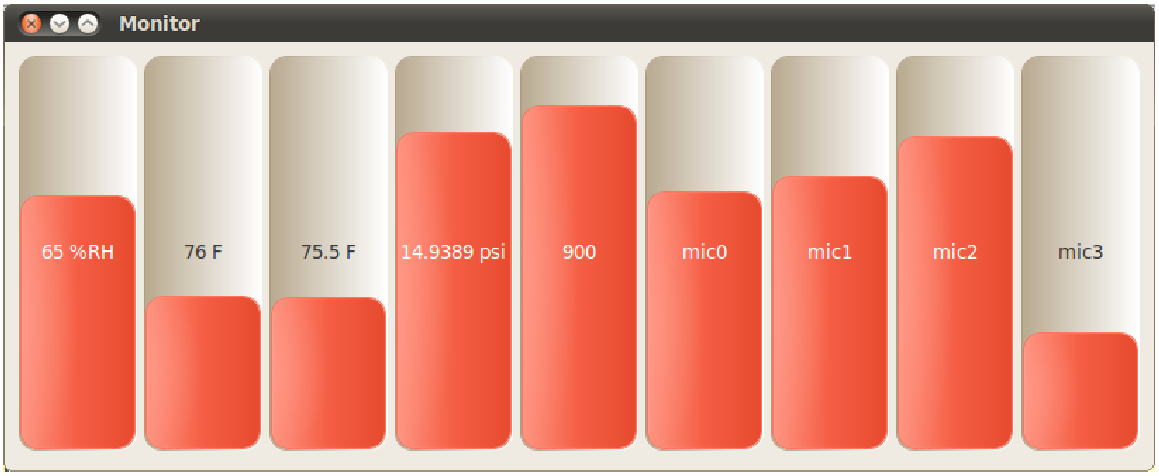

Figure 3: Screenshots of monitoring system

by heat over the course of three days for about an hour or two each day.

Chamber Tests - December 2010 Stressed was induced by both ammonia and heat, with three or four experiments per day, once a week.

Chamber Tests - February through March 2011 The experiments performed in this case were very similar to the December 2010 ones.

3.4 Noise Filtering

The experiment design was set, but one crucial concern needed to be addressed before further resources were committed to the project. This was the issue with the quality and usability of the audio recording. The presence of noise had to be evaluated and the noise itself characterized.

It can be observed in Figure 4 there is barely a noticeable difference between stressed and unstressed sound. This is due to the heavy amount of environmental noise. There are noises associated with the temperature, ammonia and humidity control, such as ventilation shafts, fans and heaters. Although these produce broadband noise, most of the noise energy is in the lower spectrum. In addition to these various sources of noise, workers routinely go into the coop to check on the birds and to

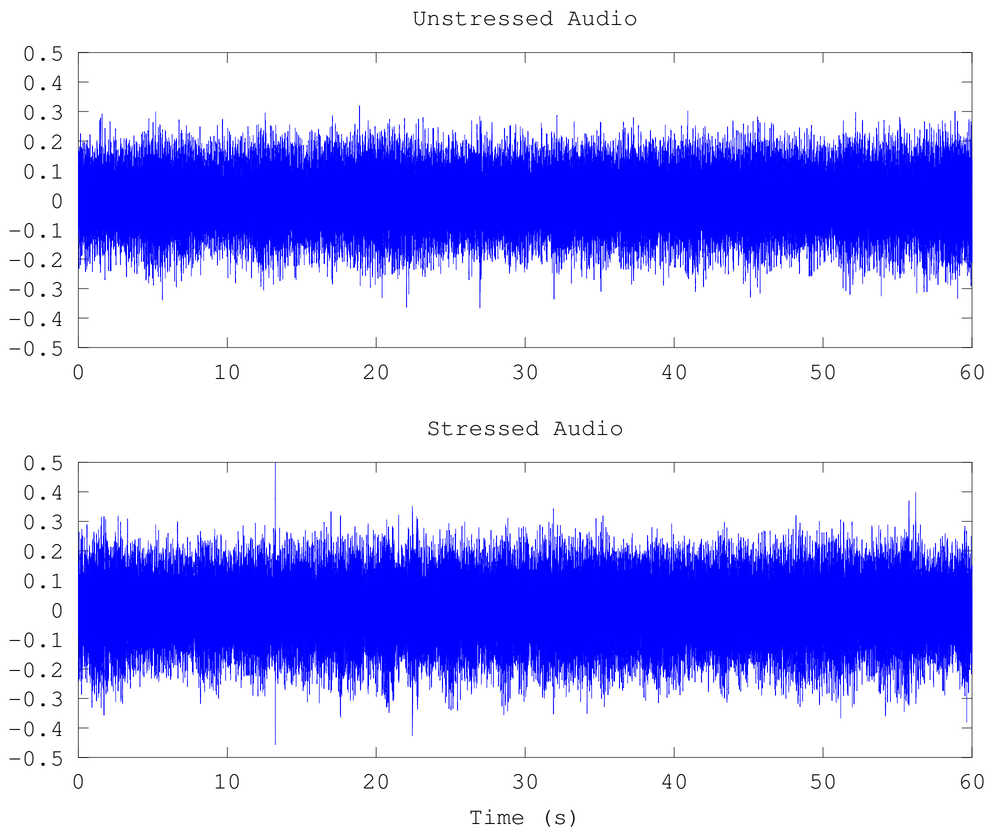

Figure 4: A Recorded Sample

remove deceased and sick birds from the flock, the noise generated by this consists, of the workers own noises (movement or speaking) and the birds feel stressed and cluck more due to the presence of potential danger. This large amount of noise means that the noise energy is much larger than the signal energy. Not only that, but the noise occurs randomly (thermostat, schedule and human intervention, among others). This prohibits a simple adaptation to the noise.

3.4.1 Noise Sources

There were several noise sources corrupting the audio:

Fans These can be a real issue. They are used to regulate temperature, and to vent ammonia. They can be turned on by sensor feedback, by schedule or by hand. They provide broadband noise going from DC to about 4 6 kHz.

Heaters These can also present problems, they are used to regulate temperature. The stress inducing tests involve heating up the chamber to an uncomfortable temperature for the chickens, an issue we confronted is that some features match perfectly with the heater noise or the noise of the fan that quickly vents out the heat as the stress cycle ends, so the classification algorithms "think" that these are the signs of stress (ignoring chicken responses).

Humans The presence of humans can be easily identified and is typically not the main source of issues.

3.4.2 Filters

Several filtering methods were tested:

Average Spectrum Error Estimate Due to the noisy characteristic of the recording it ended up behaving like a high-pass filter.

Blind Source Separation There were four microphones present for recording in the test chamber, strategically placed over the fan, over the heater and over the birds close to the feeding area. It was found that it behaved poorly, and at its best behaved like a high-pass filter.

Band-pass Filter Since the more complicated filters behaved similar to a high-pass filter, but these left some noise in the upper frequencies. It was found that a band-pass filter was better at clearing most of the noise and was much less complex. This is the filter that was used throughout the analysis.

As it can be observed in Figure 5, the details in the unstressed case can be more easily appreciated once the filter is applied. In Figure 6, the filtered stressed case can be appreciated. For clarity, Figure 7 shows both filtered outputs at an appropriate scale to observe the detail.

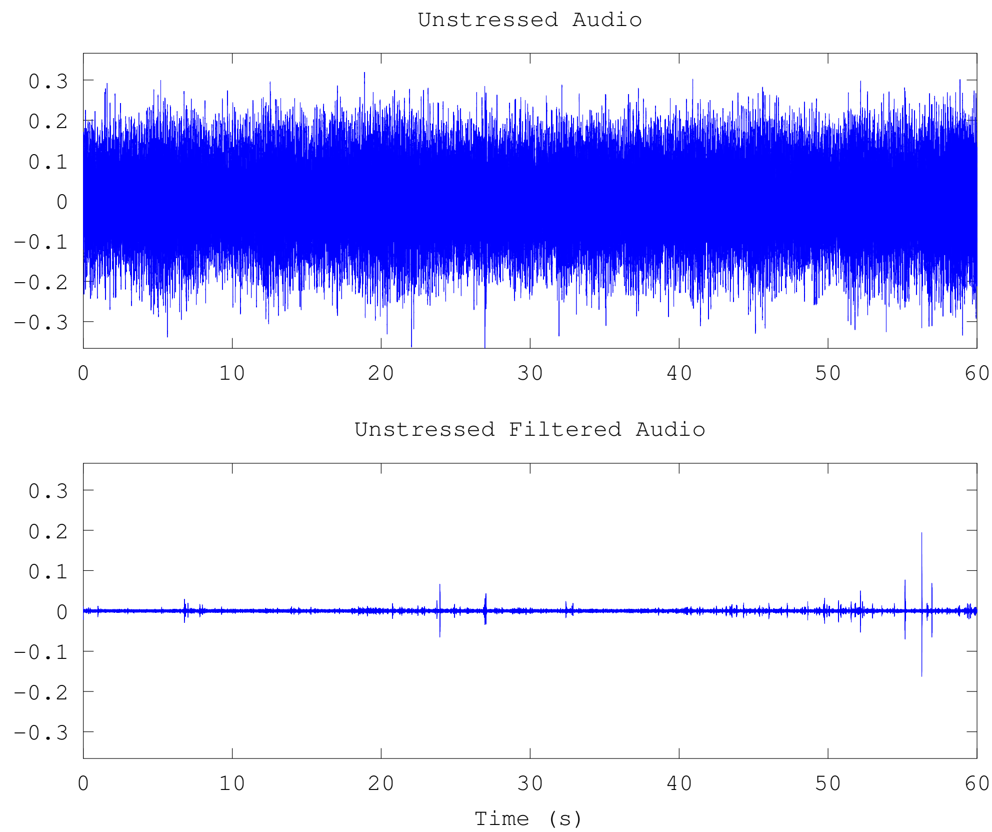

Figure 5: Band-pass Filter on Unstressed Audio

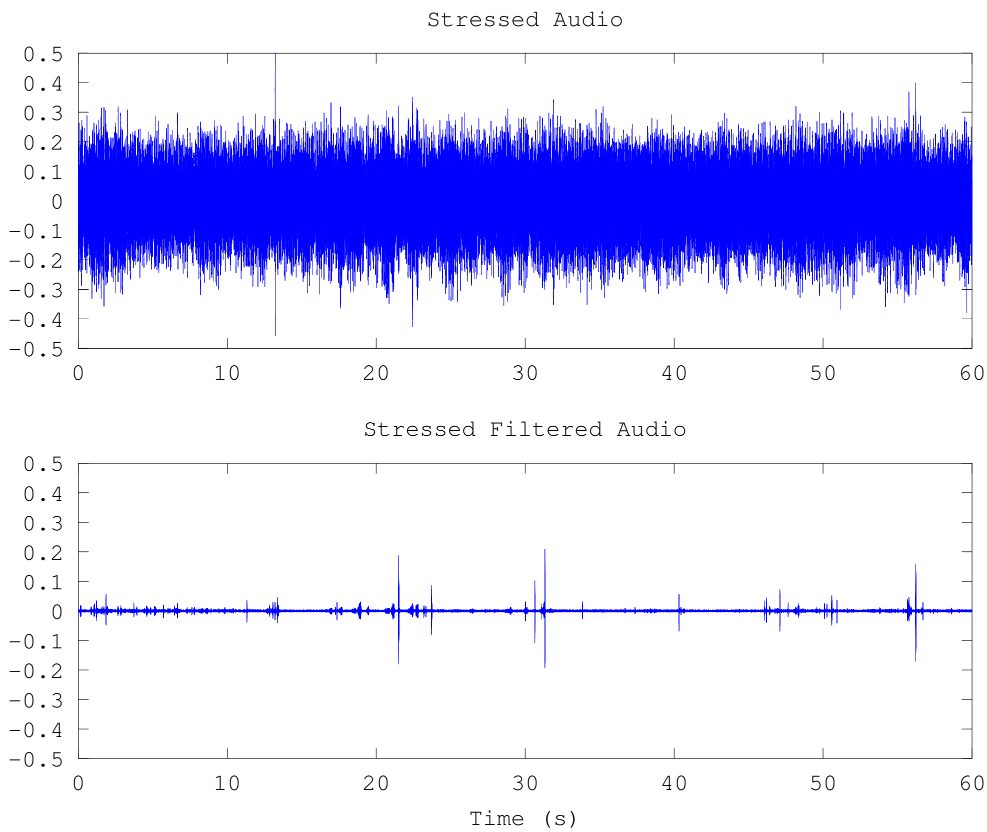

Figure 6: Band-pass Filter on Stressed Audio

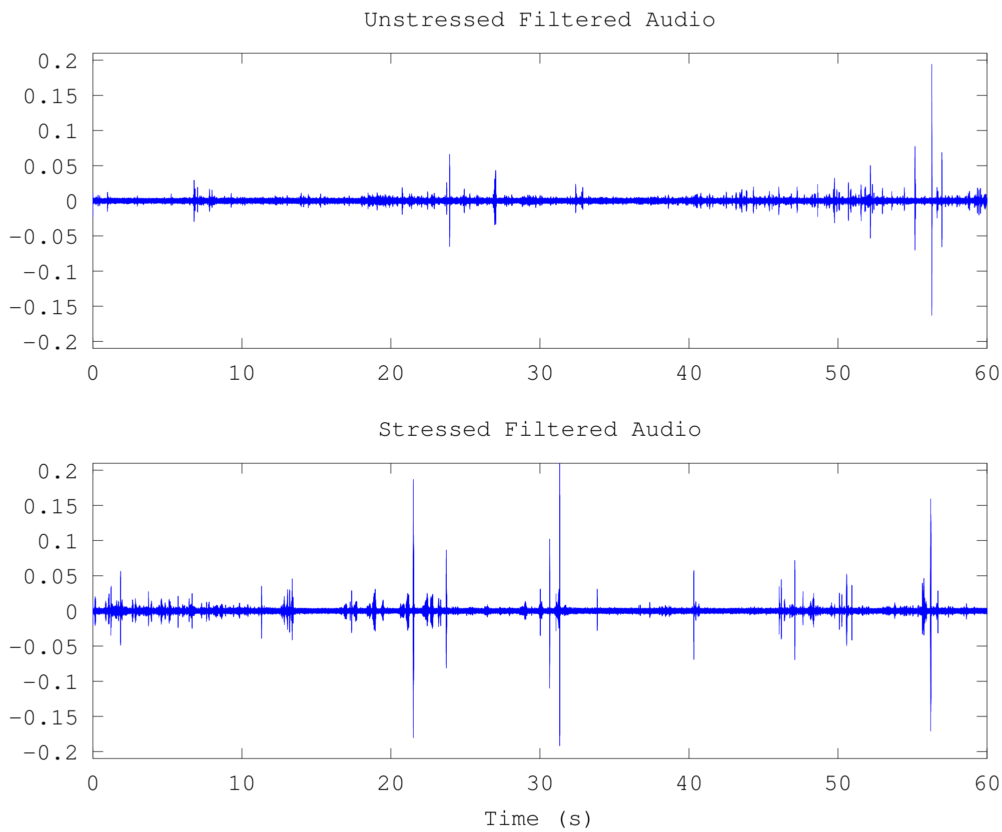

Figure 7: Band-pass Filter on Sample

CHAPTER IV

FEATURE SPACE

Typically MFCC's are used for audio analysis; however, early tests utilizing MFCC's provided poor performance for classification of these sounds. Not only this, but the MFCC's focused too much on noise sources, and were basing any classification on the characteristic noise of the heater.

4.1 Segmentation Features

Given the nature of the continuously recorded environment, good segmentation features for environmental sounds were utilized. The selection of these features was based on Gordon Wichern's continuous recording experiments. He concluded that the following features were "of specific relevance for environmental sounds." He identified three dimensions of relevance as follows:

Diversity: "the particular feature or group should exhibit a large spread (entropy) in the context of real world data sets. In particular, [functionally redundant features should be avoided] (bandwidth and spectral sparsity, for instance)."

Categoric relevance: "Different categories of sound (e.g. voice, construction sounds, traffic, wind noise) should have different average feature vectors."

Perceptual adaptation: "Sounds that sound different have different feature vectors; i.e., feature distance varies with perceptual distance. Reasonable efforts have been made to map out feature sacels according to the concept of jnd" [14].

4.1.1 Short-term features

The first set of the segmentation features consists of short-term features, which are computed every 10 ms using 20 ms of data, by means of overlapping Hamming windows.

4.1.1.1 Loudness

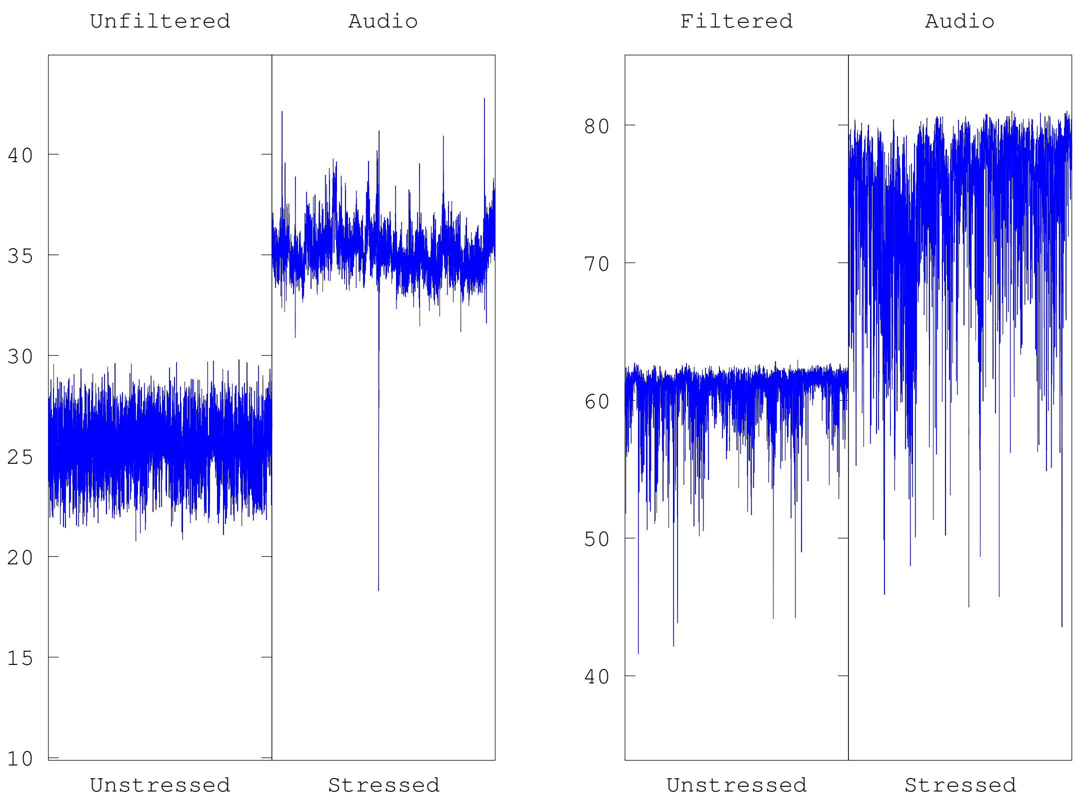

Figure 8: Loudness

The first of the short term features consists of *loudness*. It is defined as the RMS level (in decibels) of a windowed frame of data. See figure 8 for examples.

$$Y_t^{(1)} \triangleq 20 \log_{10}(\mathrm{rms}_t)$$

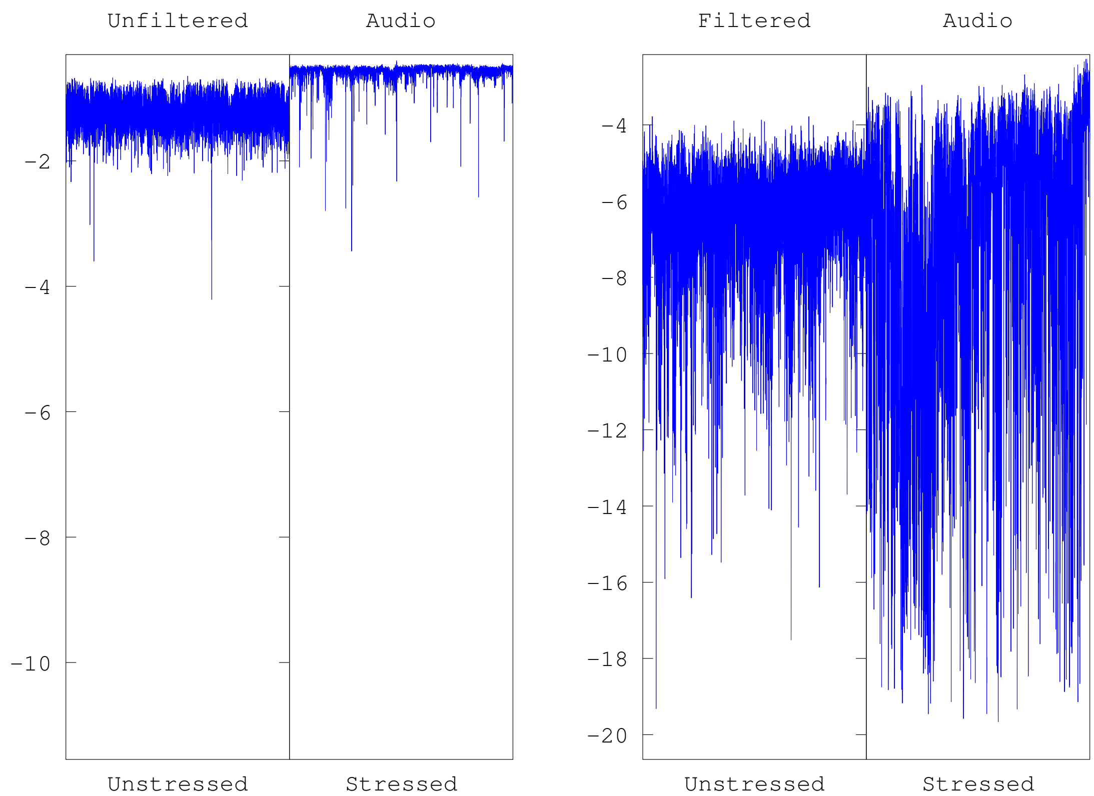

Figure 9: Spectral Centroid

4.1.1.2 Spectral Centroid

This second feature, the *spectral centroid*, is a Bark-weighted calculation of the centroid of the STFT data. See figure 9 for examples.

$$Y_t^{(2)} \triangleq \frac{\sum_{j=1}^{M} b_j \left(b_j - b_{j-1}\right) \left|X_t(j)\right|^2}{\sum_{j=1}^{M} \left(b_j - b_{j-1}\right) \left|X_t(j)\right|^2}$$

4.1.1.3 Spectral Sparsity

This feature is calculated from the zero-padded STFT data of each frame, via the ratio of L^∞ and L^1 norms calculated over the magnitude frequency spectrum. It should be large for pure tones and voice, and smaller for sounds with significant "noise" characteristic (a spread out spectrum). $X_t(j)$ is defined as the M-point

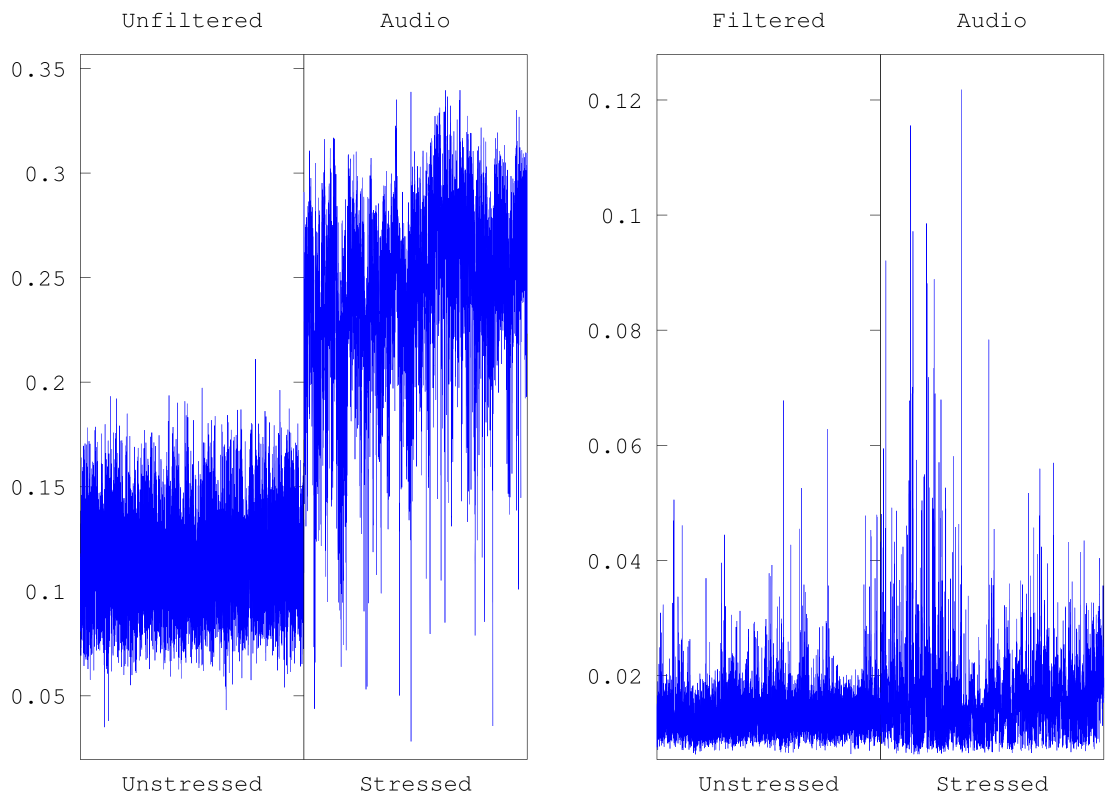

Figure 10: Spectral Sparsity

STFT coefficient from frequency bin j for frame t [14]. See figure 10 for examples.

$$Y_t^{(3)} \triangleq \frac{\max\left(|X_t(1)|, \ldots, |X_t(M)|\right)}{\sum_{j=1}^{M} |X_t(j)|}$$

4.1.2 Long-term Features

These features are calculated every 10 ms using one second worth of data by combining data from $N = 99$ of the 20 ms frames.

4.1.2.1 Temporal Sparsity

This feature is defined as the ratio of L^{∞} and L^1 norms calculated over the N small window RMS values in a given one second window. See figure 11 for examples.

$$Y_t^{(4)} \triangleq \frac{\max\left(\mathrm{rms}_{t-(N-1)}, \ldots, \mathrm{rms}_t\right)}{\sum_{k=t-(N+1)}^{t} \mathrm{rms}_k}$$

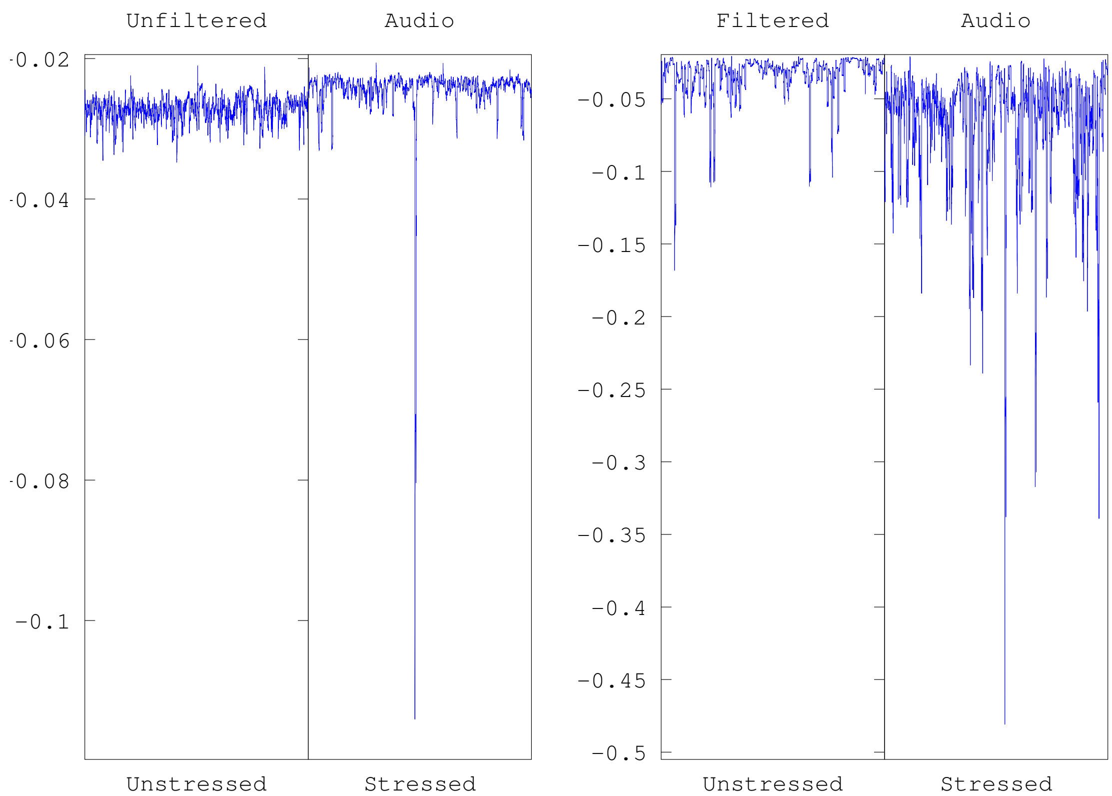

Figure 11: Temporal Sparsity

4.1.2.2 Transient Index

This feature is computed by combining the Mel frequency cepstral coefficients from several frames of data. See figures 12 for examples.

$$Y_t^{(5)} \triangleq \sum_{k=t-(N+2)}^{t} \|\mathrm{MFCC}_k - \mathrm{MFCC}_{k-1}\|_2$$

4.1.2.3 Harmonicity

This feature is used to measure probabilistically whether or not the STFT spectrum for a given frame exhibits a harmonic frequency structure. The value of this feature is high for speech, music, and machine and low for environmental audio and bells. This is a useful number, because hopefully it will be robust to noise. See figure 13

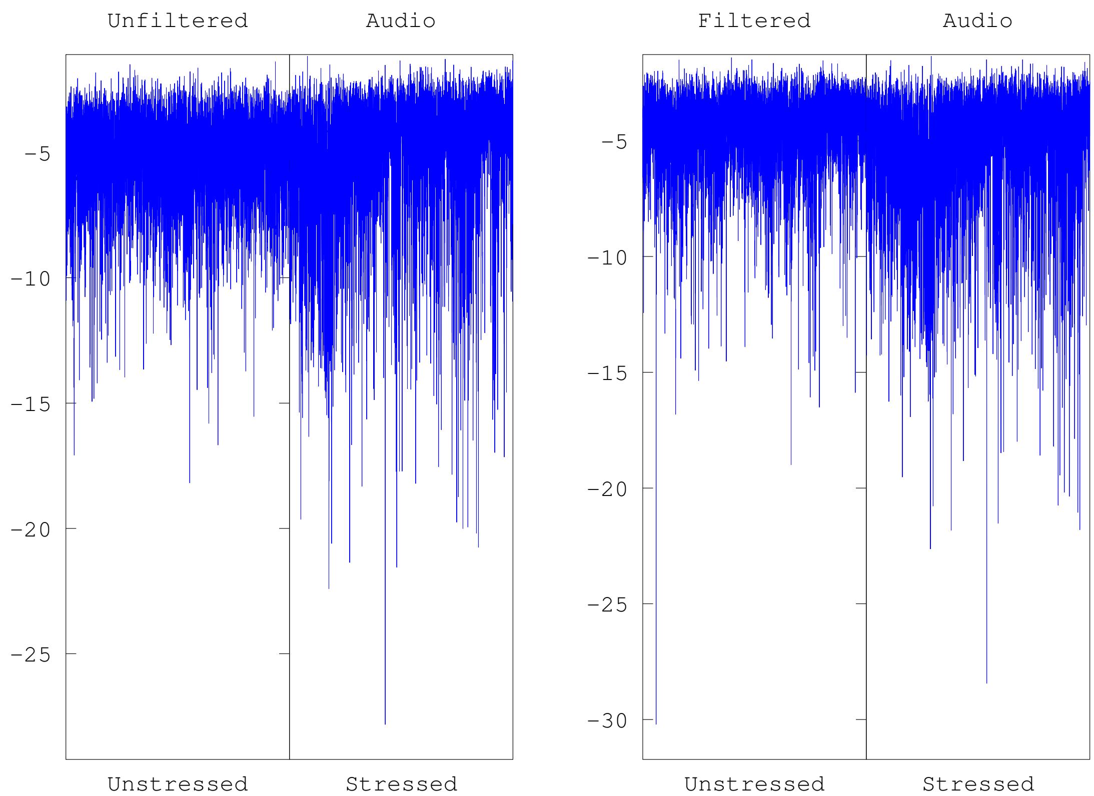

Figure 12: Transient Index

for examples.

$$
\begin{aligned}
Y_t^{(6)} &\triangleq \max_{f_1,f_2,k_1,k_2} P\left(f_1, f_2|\hat{f}_0, k_1, k_2\right) \\
&\text{s.t.} : \{f_1, f_2\} \subset \rho_t, 1 \leq k_1 < k_2 \leq k_{\max}, f_1 < f_2, \text{and} \hat{f}_0 > f_{\min}, \\
\hat{f}_0 &= \frac{(f_1/k_1)^2 + (f_2/k_2)^2}{f_1/k_1 + f_2/k_2} \\
\sigma_j &= C k_j f_0 \\
P\left(f_1, f_2|\hat{f}_0, k_1, k_2\right) &= \prod_{j=1}^{2} \gamma\left(f_j; k_j f_0, \sigma_j\right) \\
C &= 0.01/\sqrt{2}
\end{aligned}
$$

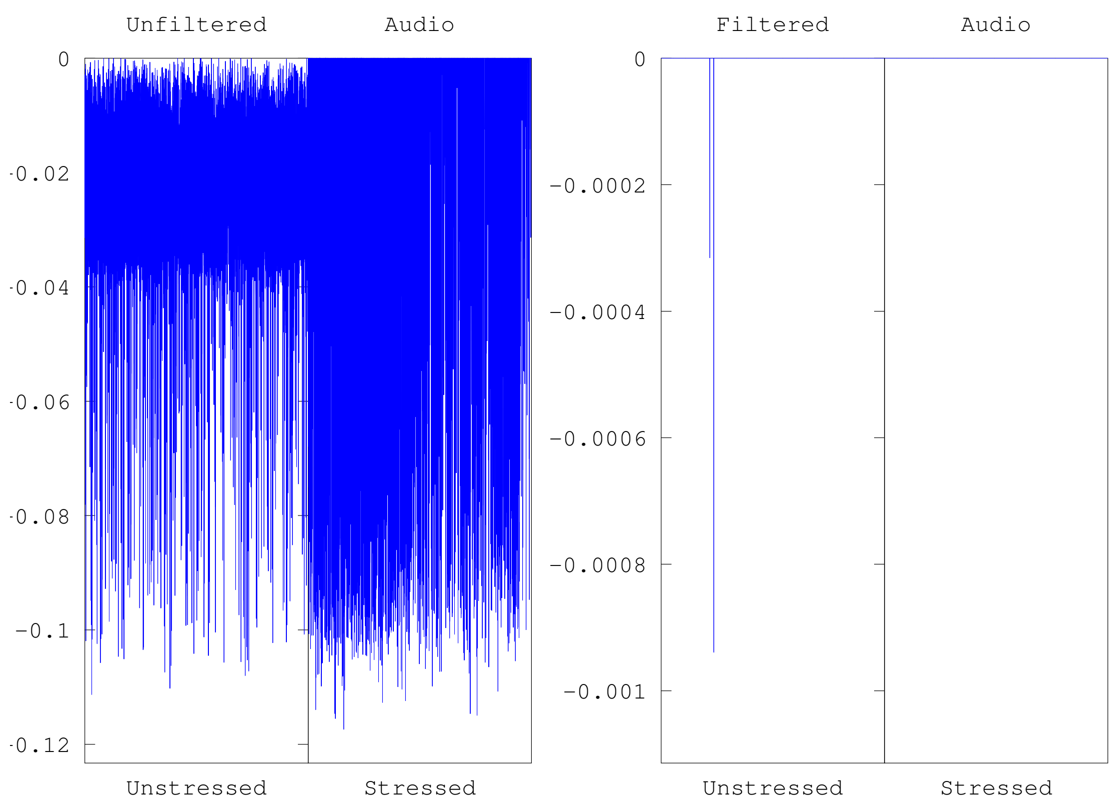

Figure 13: Harmonicity

4.2 First Four Temporal Central Moments

When going through the data acquired, a certain "peakedness" was noticed on the audio whenever the birds where stressed. Due to the notable difference in the temporal data, the temporal central moments where evaluated as part of the feature space. Of special interest was the kurtosis, which is traditionally a measure of "peakedness." Since the kurtosis was being computed the other central moments where also considered. These were all computed over one second rectangular non-overlapping windows over the temporal data.

4.2.1 Mean

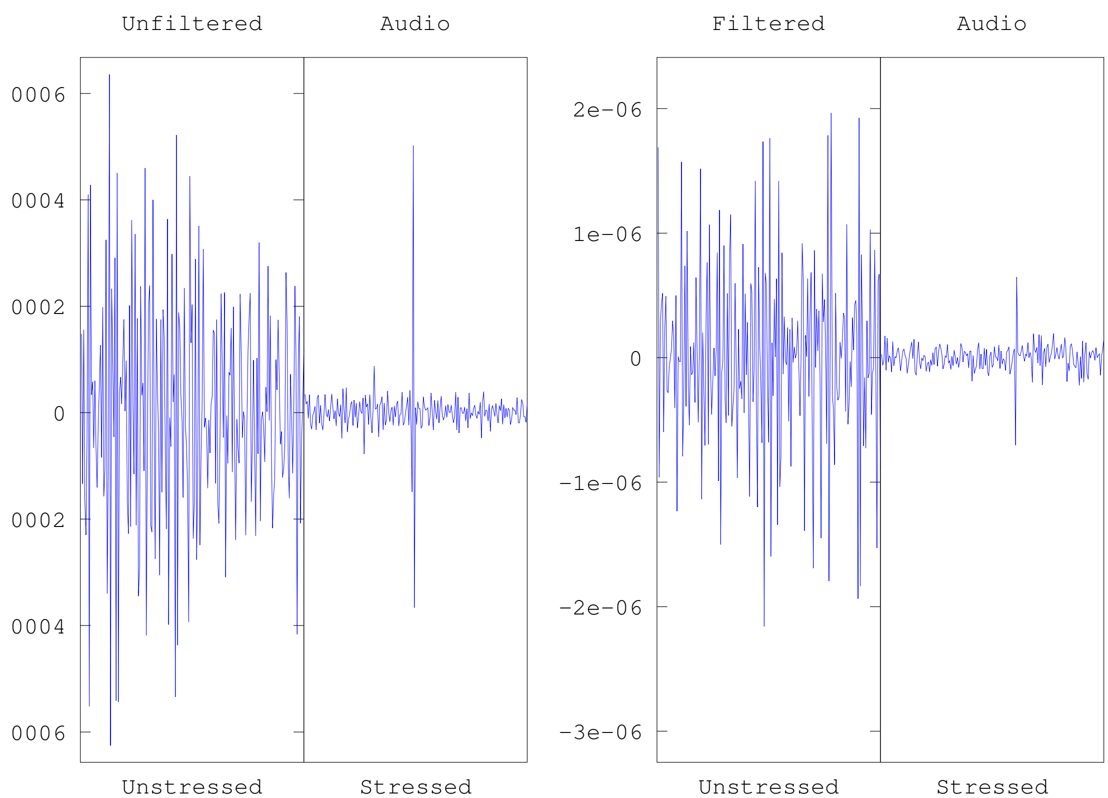

Figure 14: Mean

Technically speaking this is not a central moment, the first central moment is always zero. This is the first moment, and although the it should be about zero (the alternating nature of audio waves), it was taken into consideration to study the presence of any DC components in the originating signal. See figure 14 for examples.

$$\mu = E[X]$$

4.2.2 Standard Deviation

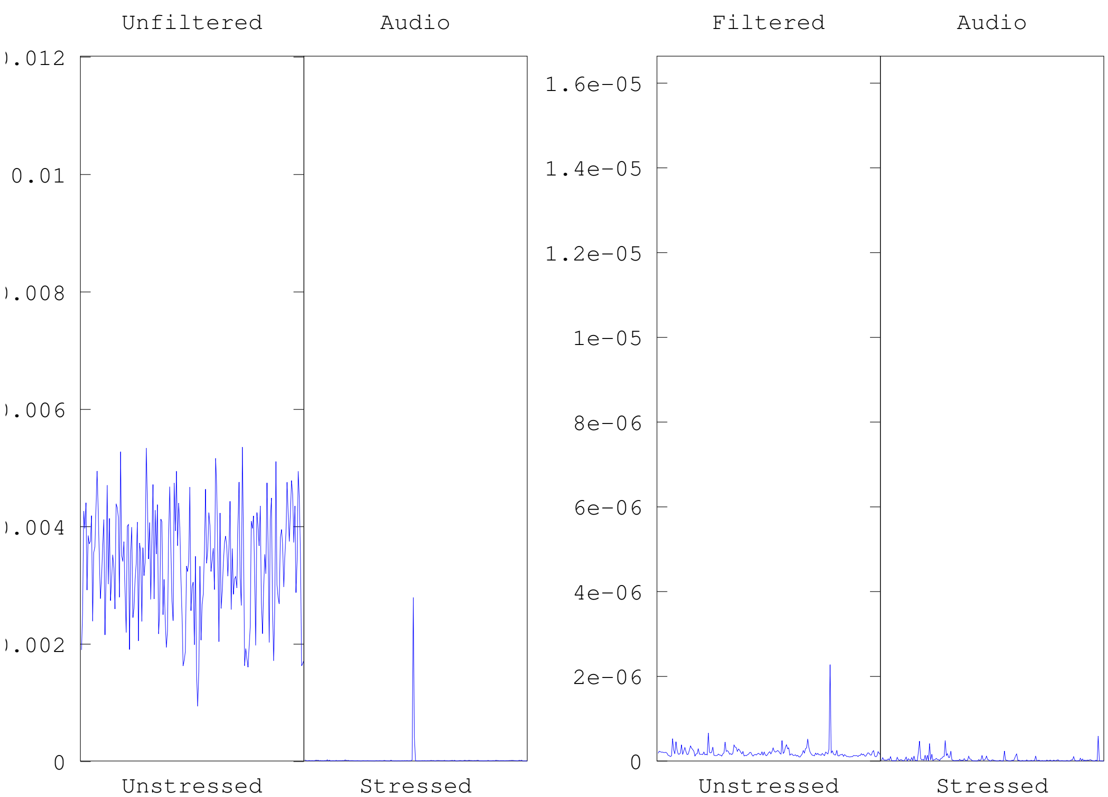

Figure 15: Standard Deviation

The standard deviation estimates the amplitude of the signal. See figure 15 for examples.

$$\sigma = \sqrt{E\left[(X-\mu)^2\right]}$$

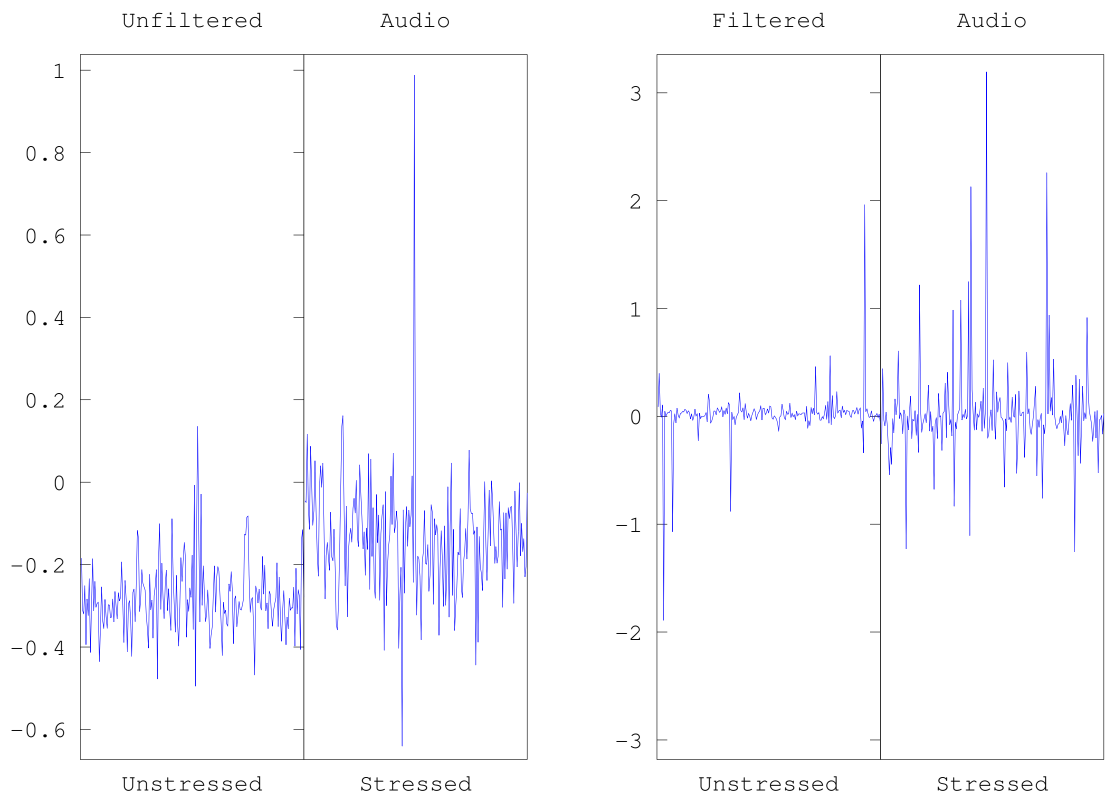

Figure 16: Skewness

4.2.3 Skewness

The skewness serves as a measure of the "lopsideness" of the probability distribution behind the data. See figure 16 for examples.

$$\gamma_1 = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$$

4.2.4 Kurtosis

As mentioned earlier, this is the most interesting statistical measure. It measures the "peakedness," which can be used to determine when the birds are making a ruffle, meaning whenever their excited or scared. It should present some noise robustness as the noise typically belongs to a somewhat Gaussian distribution. See figure 17 for

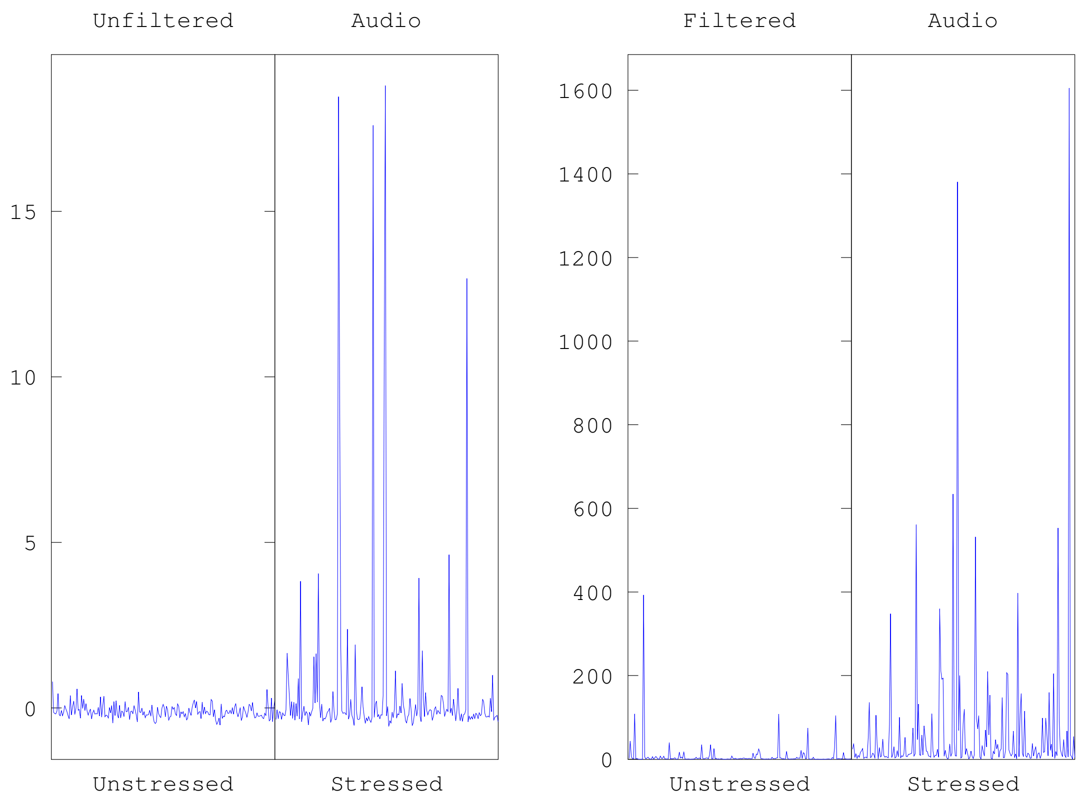

Figure 17: Kurtosis

examples.

$$\gamma_2 = \frac{E\left[(X-\mu)^4\right]}{E\left[(X-\mu)^2\right]^2}$$

4.3 Promising Features

From what can initially be observed, kurtosis shows excellent promise as a feature for classification, and it is fairly easily computed. This would make it the best feature if it continues to show such promise in further analysis. Other features such as loudness show good promise.

In order to better evaluate these features, simple smoothing filters were applied to them post calculation. There was a simple smoother and a more aggressive smoother, examples of each on the kurtosis feature can be observed in figures 18 and 19. In the next chapter, these smoothed results will be used to analyze the features. The smoothing was performed with a low-pass with the cut-off at .01 and .001 (aggressive).

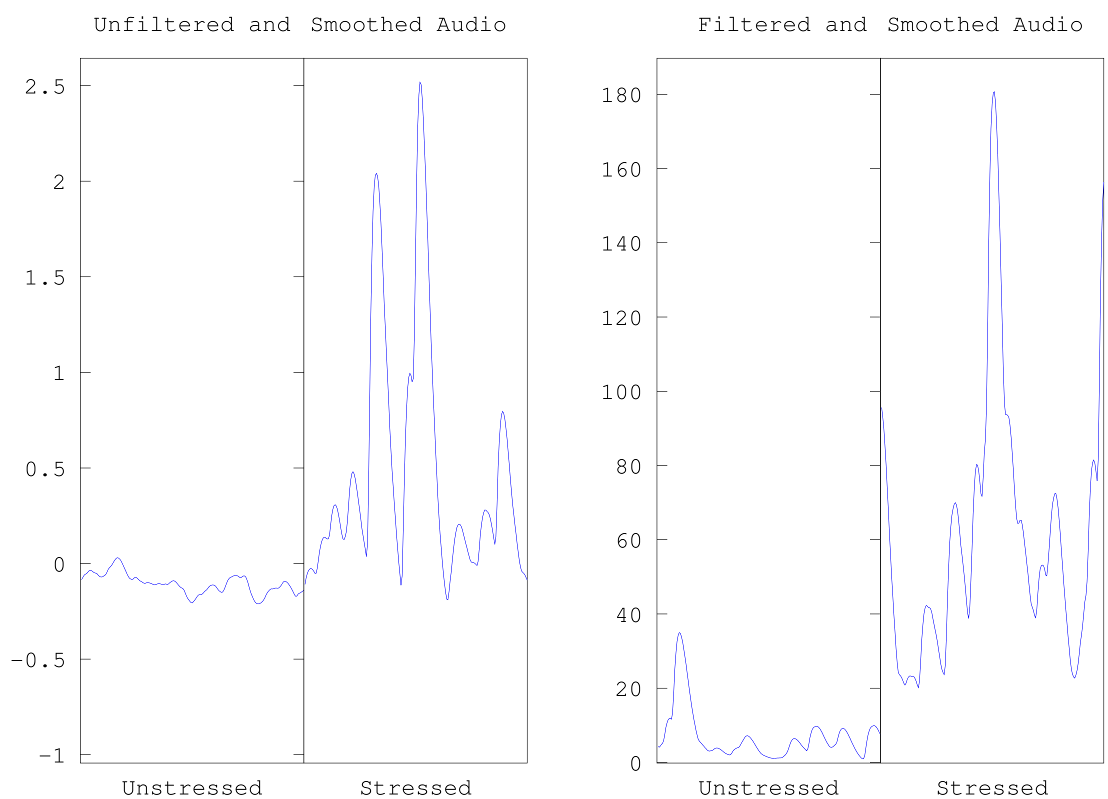

Figure 18: Smoothed Kurtosis

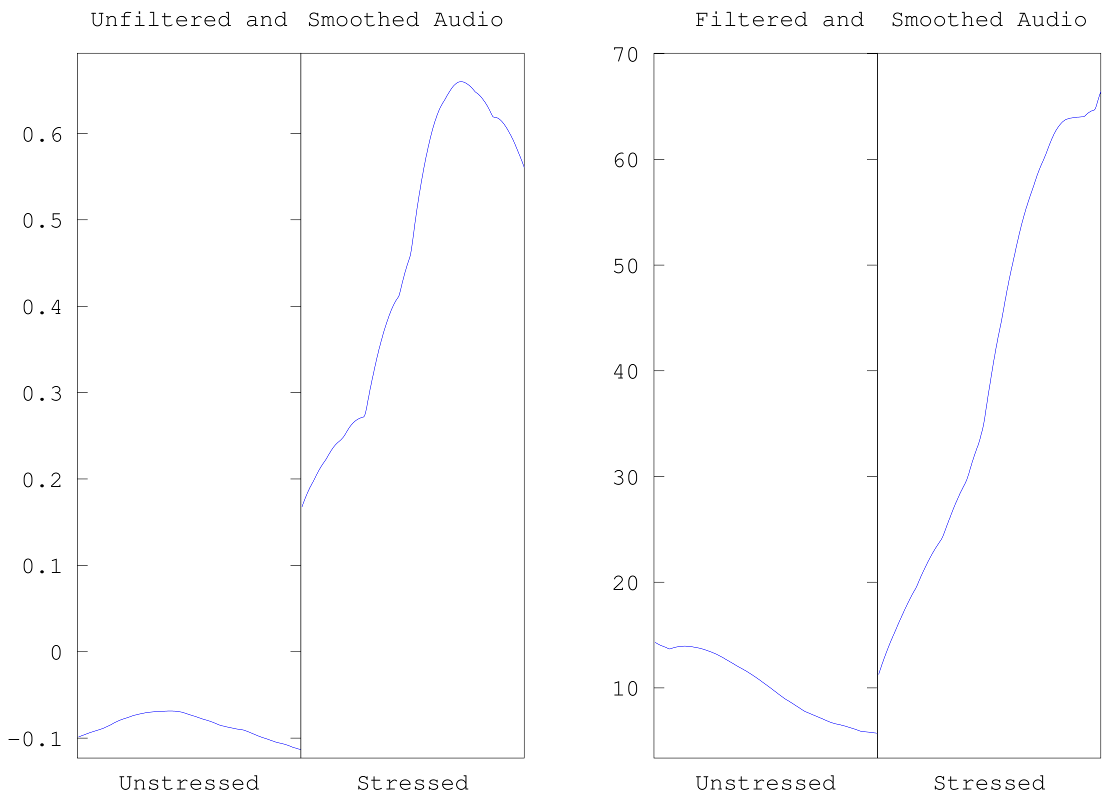

Figure 19: Aggressively Smoothed Kurtosis

CHAPTER V

FEATURE EVALUATION AND ANALYSIS

5.1 Desired Information

As mentioned earlier the goal is to be able to determine whether a bird is stressed or not by means of its vocalizations. Instead, increased heat and ammonia levels were used to induce stress, and then audio recordings were used to detect when this induced stress occured.

The features were individually evaluated for their performance in detecting stress based on a simple threshold classifier. The threshold was swept through all possible values, and the classifier's performance was measured and ROC curves and accuracy were generated. The usefulness of the features themselves was determined from the ROC curves. This individual attention to detail is necessary to better understand how the features are behaving. The features were also collectively evaluated by means of an Adaboost classifier. Its performance was measured by means of the area under ROC curves and the accuracy.

5.1.1 Simple Threshold Classifier

5.1.1.1 ROC curves

An ROC curve is a plot of the true positive rate (recall) versus the false positive rate (fall-out) of a classifier. It is built by computing the recall and fall-out at each threshold setting. This allows for selecting the most optimal feature. In these curves, the diagonal dotted line represents a classifier that is correct 50% of the time (random classifier). A good feature is one that is as close as possible to the axes on the upper left hand side of the diagonal (or have an area of 1 under the curve). ROC curves were generated for each feature.

Another thing that can be evaluated using these plots is robustness. If the raw signal and the filtered signal have similar outcomes (shapes), then it can be somewhat assumed that the feature is robust to noise (i.e., it behaves the same with or without noise).

Finally, the legend needs some explanation. The argument on the right-hand side explains if the audio signal has been filtered or not, and the argument on the left hand signal refers to whether the extracted feature has been smoothed or not. There are six lines in each plot:

Raw on raw. The feature is extracted from the original signal and it does not undergo any further processing.

Smoothed on raw. The feature is extracted from the original signal and it is smoothed by a low-pass filter.

Aggressively smoothed on raw. The feature is extracted from the original signal and it is smoothed by a more aggressive low-pass filter.

Raw on band-passed. The feature is extracted from a filtered (band-pass filter) version of the audio signal and it does not undergo any further processing.

Smoothed on band-passed. The feature is extracted from a filtered (band-pass filter) version of the original signal and it is smoothed by a low-pass filter.

Aggressively smoothed on band-passed. The feature is extracted from a filtered (band-pass filter) version of the original signal and it is smoothed by a more aggressive low-pass filter.

5.1.1.2 Analysis

This first set of ROC curves are for the 12/03/2010 experiments. These will be analyzed first, the best features will go on to the next analysis.

Figure 20: Loudness. This feature seems to be fairly robust, both the original (Raw on raw) and the filtered (Raw on band-passed) seem to show similar results (the band-passed performed slightly better). Anything else (smoothing) did not affect the results much at all (for the better or for the worse). All curves are above the "chance" line and are fairly close to the axes. This is a good feature for heat stress. When it comes to classifying ammonia stress it behaved more like chance.

Figure 21: Spectral Centroid. This seems to be a classic example of the feature identifying the heater or the fan and not the birds. This conclusion comes from observing how well the unfiltered version fairs against the filtered version of the signal. The filtered version behaves like a random classifier. Once the heater noise is removed the classifer does not work. This is a very poor feature. In the ammonia case it performed poorly.

Figure 22: Spectral Sparsity. This feature seems to suffer from a similar problem as the Spectral Centroid. It seems to be keying on the noise from the heater/fan and not on the birds themselves. Poor performance in the ammonia case

Figure 23: Temporal Sparsity. This feature also seems to behave like chance.

Figure 24: Transient Index. This feature also seems to behave like chance.

Figure 25: Harmonicity. This feature suffers from apparently identifying itself with the heater and such, as we can see good performance in the unfiltered case and poor performance otherwise.

Figure 26: Average. This feature also seems to behave like chance.

Figure 27: Standard Deviation. This is poor feature and behaves like chance in the ammonia case.

Figure 28: Skewness. This feature suffers from apparently identifying itself with the heater and such, as we can see good performance in the unfiltered case and poor performance otherwise.

Figure 29: Kurtosis. This feature had examplary behavior. The feature on both the filtered and the original audio behaved extremely similarly, making it robust. Second, with each smoothing it behaved more and more like a perfect classifier. In the case of identifying ammonia stress it didn't do as well.

It has been observed that in the 12/03/2010 experiment, the best features were the kurtosis and the loudness. Now that the best features in that experiment have been defined, their performance will be corroborated in other experiments. It was apparently very hard to identify ammonia stress, this is due to the fact that in this experiment the ammonia stress tests were introduced for the first time and stress was not induced that well. The performance seems to improve in other experiments.

In the case of loudness, in Figure **??** it can be observed that it continues to be a good feature for identifying heat induced stress, but poor at identifying the ammonia induced stress. In Figure **??** we can observe that the feature performed extremely well in both cases. Finally in Figure **??** the performance appeared to be good but not great.

Kurtosis performed very well in both heat and ammonia stress in Figure **??**, but smoothing yielded some strange results. In Figures **??** and **??** there continues to be great performance except with the problem pertaining to smoothing.

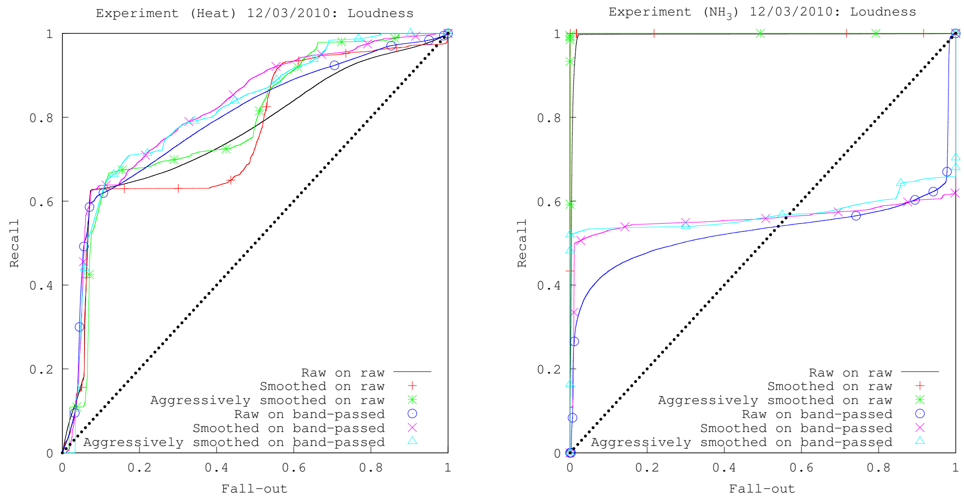

Figure 20: ROC on 12/03/2010 on the Loudness

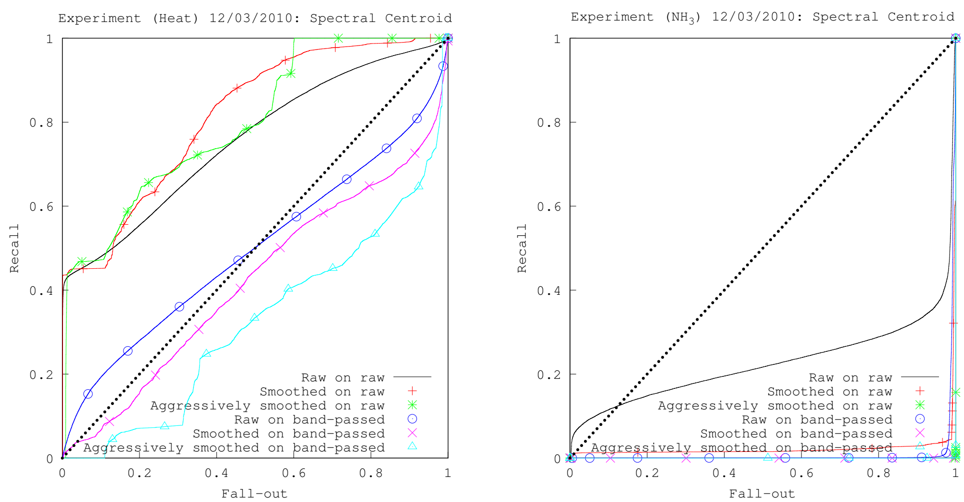

Figure 21: ROC on 12/03/2010 on the Spectral Centroid

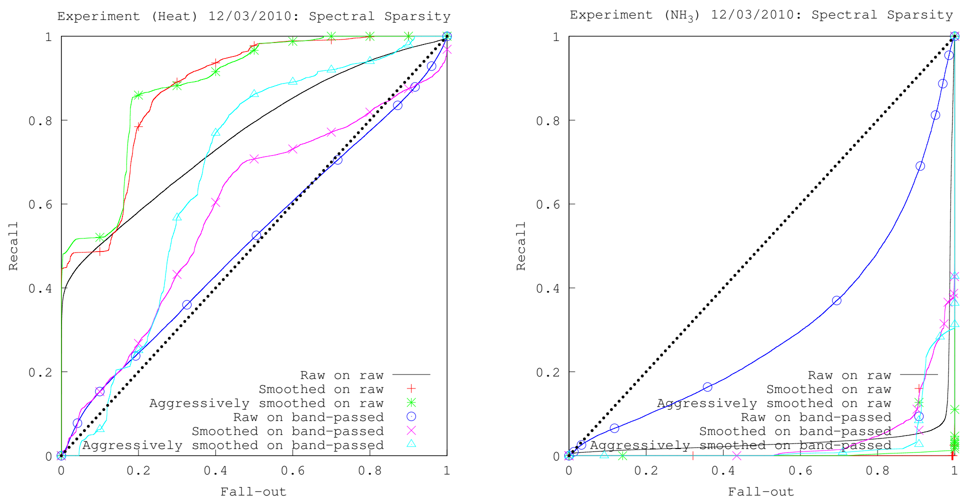

Figure 22: ROC on 12/03/2010 on the Spectral Sparsity

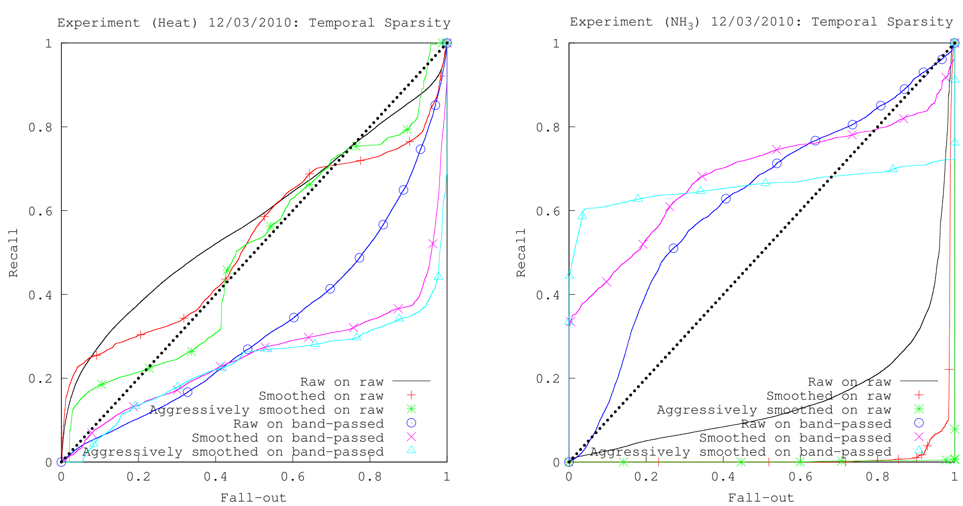

Figure 23: ROC on 12/03/2010 on the Temporal Sparsity

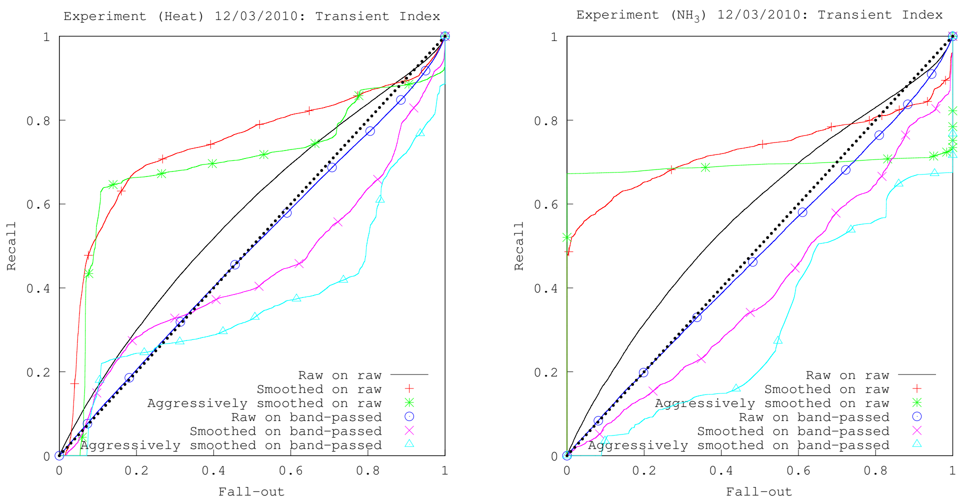

Figure 24: ROC on 12/03/2010 on the Transient Index

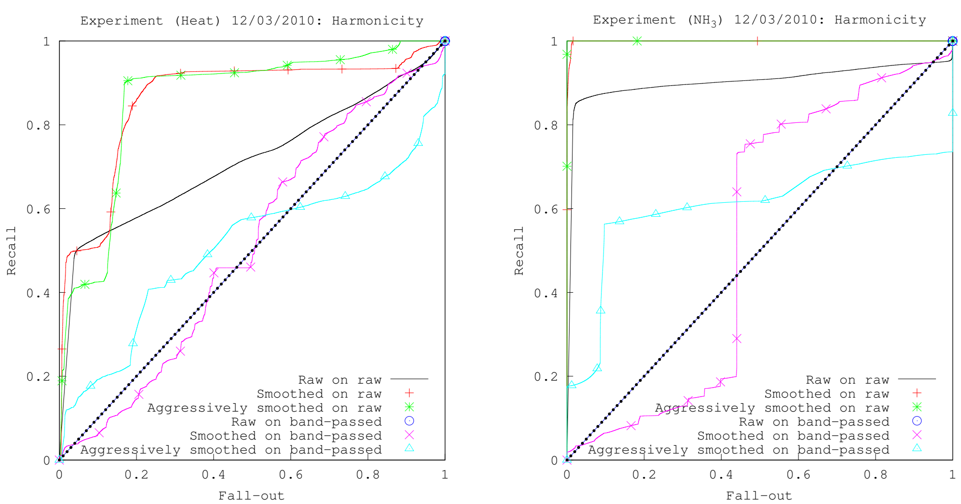

Figure 25: ROC on 12/03/2010 on the Harmonicity

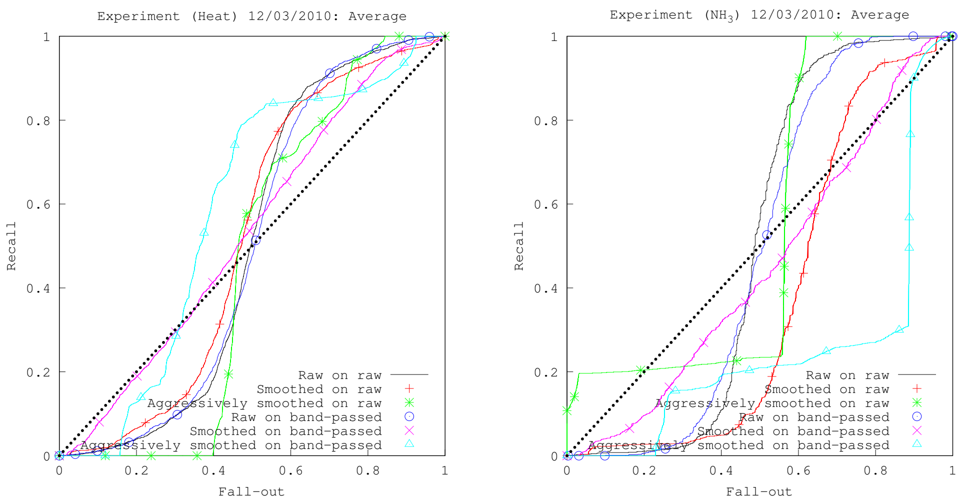

Figure 26: ROC on 12/03/2010 on the Average

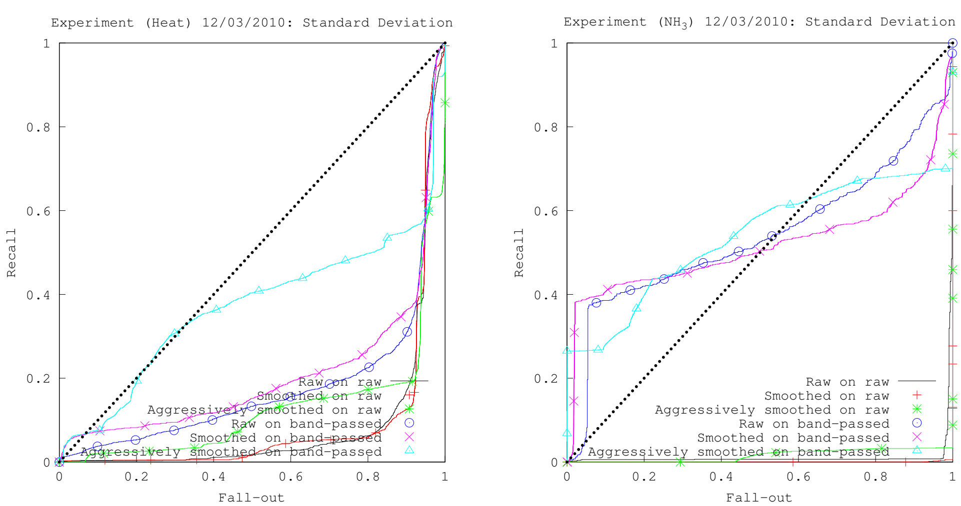

Figure 27: ROC on 12/03/2010 on the Standard Deviation

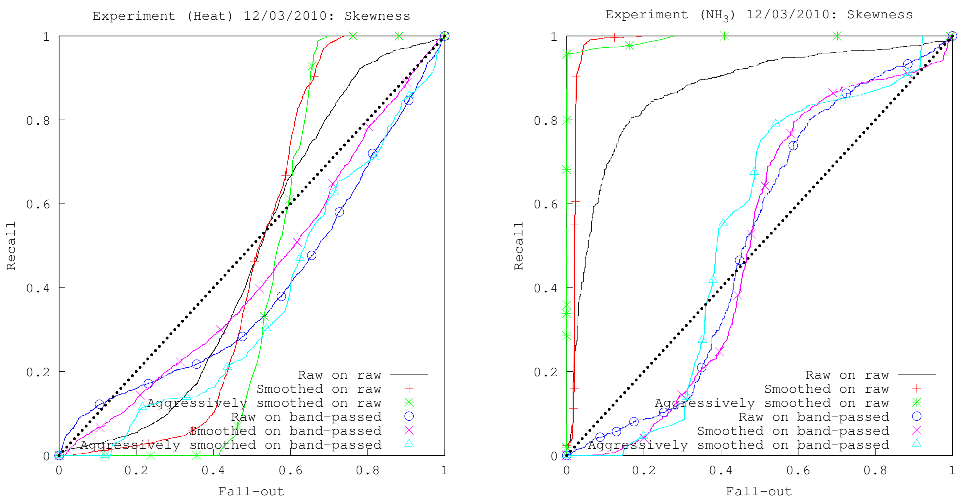

Figure 28: ROC on 12/03/2010 on the Skewness

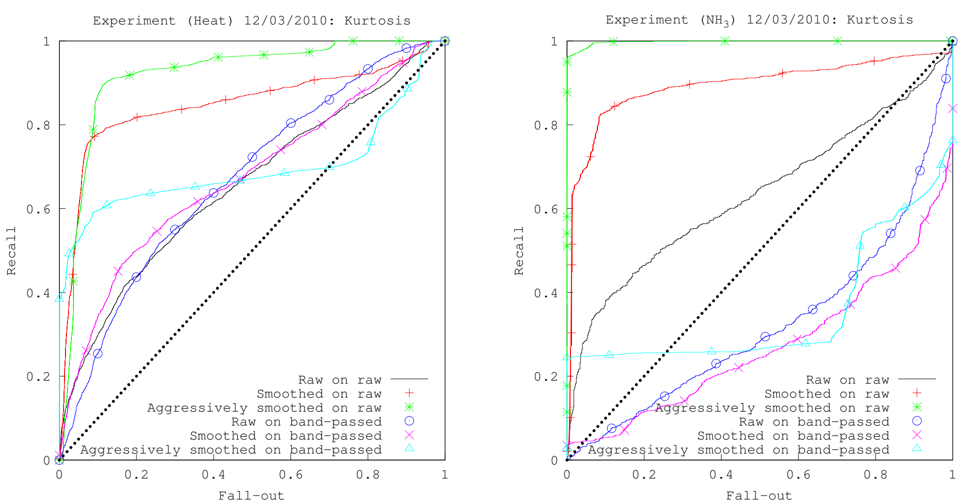

Figure 29: ROC on 12/03/2010 on the Kurtosis

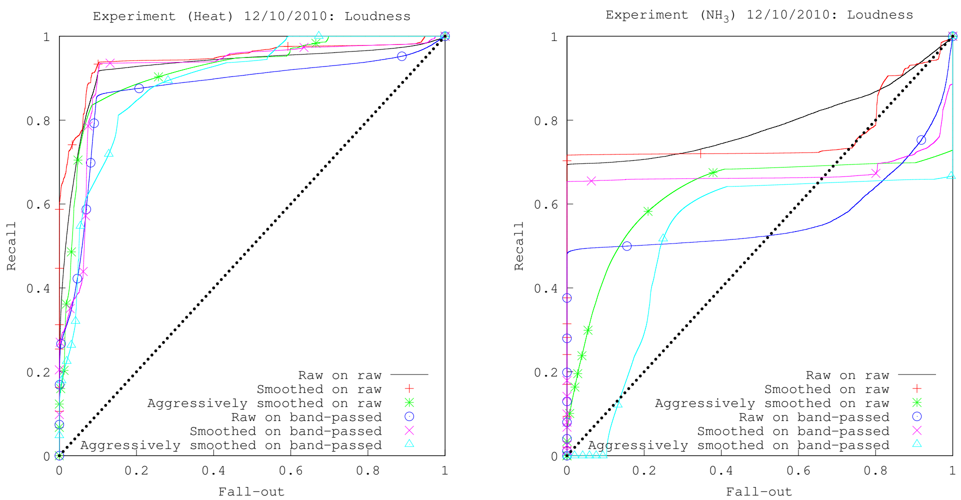

Figure 30: ROC on 12/10/2010 on the Loudness

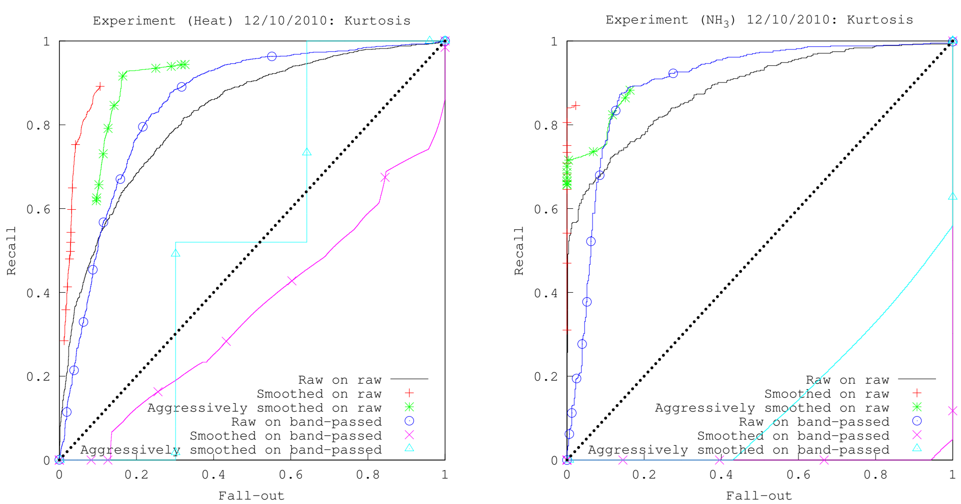

Figure 31: ROC on 12/10/2010 on the Kurtosis

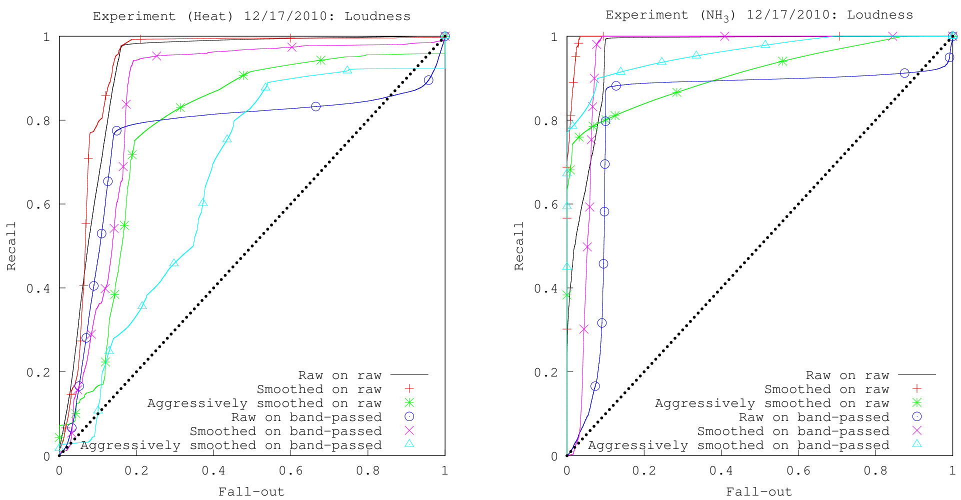

Figure 32: ROC on 12/17/2010 on the Loudness

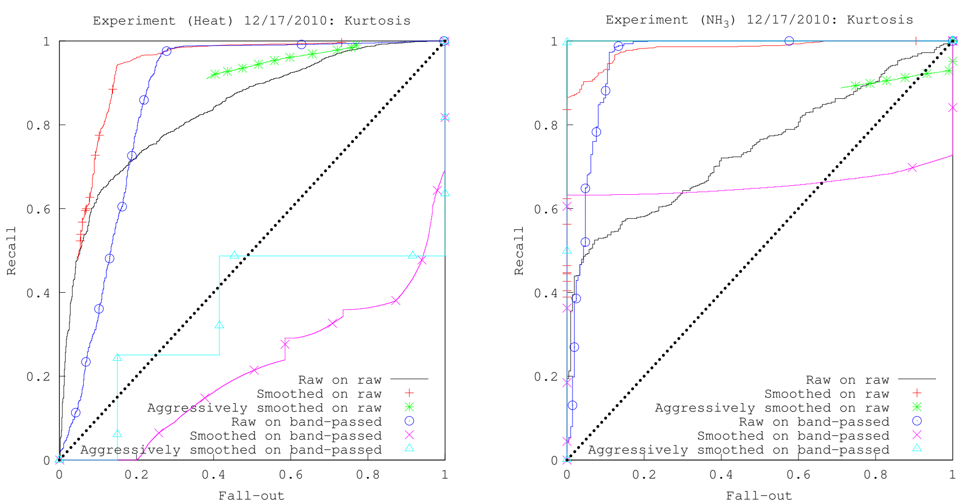

Figure 33: ROC on 12/17/2010 on the Kurtosis

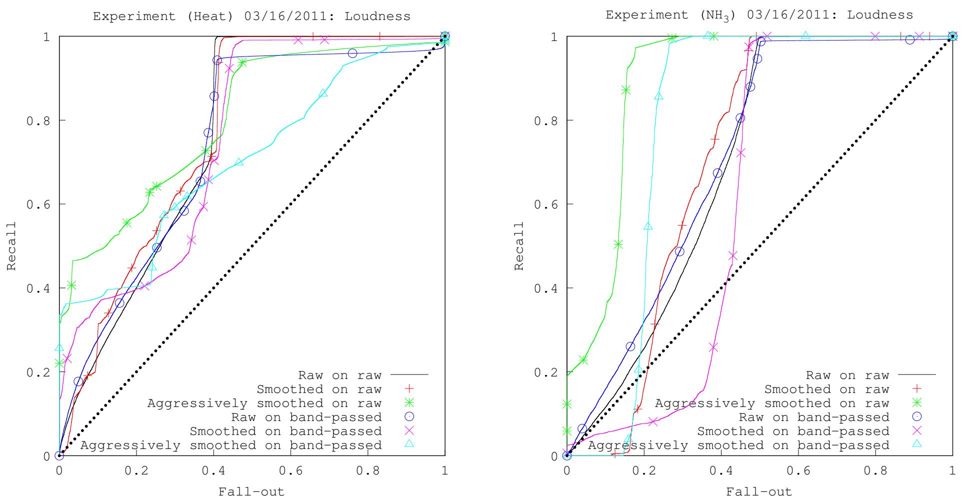

Figure 34: ROC on 03/16/2011 on the Loudness

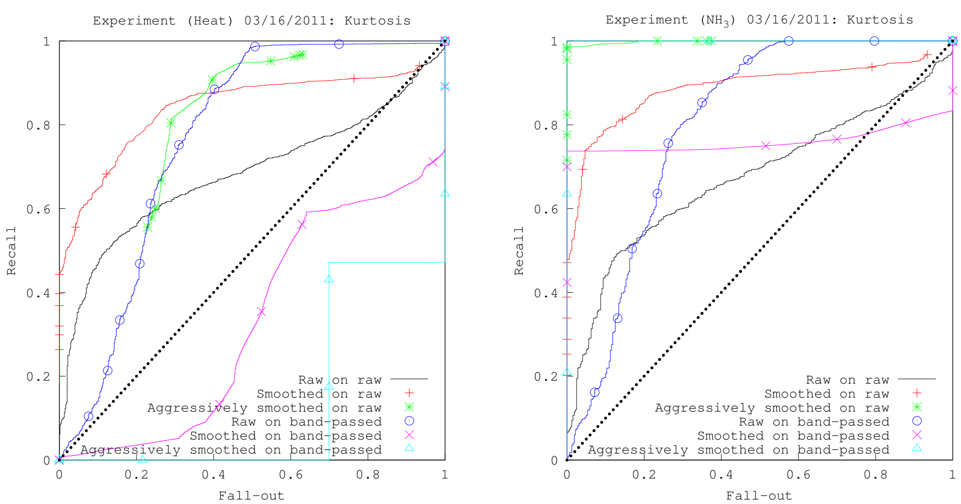

Figure 35: ROC on 03/16/2011 on the Kurtosis

Table 1: Maximum Accuracy on Heat Experiment 12/03/2010.

Feature	Orig	F1	F2	BP	BP F1	BP F2
Avg	0.84007	0.84007	0.84007	0.84082	0.84082	0.84082
Std Dev	0.84007	0.84007	0.84007	0.84082	0.84082	0.84082
Skewness	0.84065	0.84007	0.84007	0.84082	0.84082	0.84082
Kurtosis	0.84007	0.84007	0.84007	0.84082	0.84082	0.84082
Loudness	0.87868	0.87810	0.86017	0.87535	0.87987	0.86708
Spec Centroid	0.90250	0.90867	0.89942	0.84076	0.84076	0.84076
Spec Sparsity	0.89649	0.91115	0.91348	0.84076	0.84076	0.84076
Temp Sparsity	0.84641	0.85463	0.84001	0.84076	0.84076	0.84076
Transient Idx	0.84001	0.85555	0.85063	0.84076	0.84076	0.84076
Harmonicity	0.88501	0.90125	0.88234	0.84076	0.84208	0.84551

Table 2: Maximum Accuracy on Heat Experiment 12/03/2010.

Experiment	Type	Kurtosis	Loudness
12/03/2010	Heat	0.84007	0.87868
12/03/2010	NH3	0.74301	0.99027
12/10/2010	Heat	0.74770	0.90699
12/10/2010	NH3	0.79961	0.83739
12/17/2010	Heat	0.78531	0.93872
12/17/2010	NH3	0.79961	0.97253
03/16/2011	Heat	0.85439	0.93948
03/16/2011	NH3	0.85708	0.92741

5.1.1.3 Accuracy

One can appreciate the power of ROC curves as a good tool to identify strong classifier, from observing the contrast with the behaviour of maximum accuracy. According to Table 1 the spectral centroid is the best feature for identifying heat stress, but it can be clearly observed from the ROC curve that it was not. Over the 12/03/2010 experiment kurtosis achieved a maximum accuracy of 84.0% and loudness 87.9%, in Table 2 some other results can be appreciated, for the unfiltered unsmoothed cases. In these simple cases loudness showed to have great accuracy in distinguishing between stress and no stress, and kurtosis having some fare results.

5.1.2 AdaBoost Classifier

The reason for this was to observe the combined behavior of the features. The classifier was built using the AdaBoost M1 algorithm with 10 iterations, 1 seed, no resampling and a weight threshold of 100 and using a decision stump (using entropy) as the weak classifier. All of the results shown here are based on a 10-fold cross-validation. The data with regards to NH_3 on the 12/03/2010 was not included in this analysis, due to its related acquisition problems. The data over four experiments was combined into this analysis. All of the segmentation features were downsampled (they were computed at a higher rate) to match the 1Hz rate of the statistical features. When building a model using all of the combined heat data (19783 instances) the stratified accuracy was 95.0412% with an area under the ROC being 0.987. When building a model using all of the combined NH_3 data (4195 instances) the stratified accuracy was 97.8784% with an area under the ROC being 0.999. When building a model combining both stress inducers (23978 instances) the stratified accuracy was 94.2280% with an area under the ROC being 0.986. Now these were built using all 60 dimensions (the 10 features, over original and filtered audio, not-smoothed, smoothed and aggressively smoothed). The biggest problem this has is all those features that focused on the noise of the heater are playing in and not helping. Using only the filtered audio data (which makes more sense, because it eliminates contributions due to the heater and other sources of noise), the achieved accuracy was 91.8509% and the area under the ROC curve was 0.976, which is still a very good classifier. Using only the loudness and kurtosis on the unfiltered audio data the classifier it achieved an accuracy of 91.684% and an area under the ROC curve of 0.973, when using these features on the filtered audio the accuracy was 90.191% and the area under the ROC curve was 0.960. Using only the kurtosis on the filtered audio the accuracy was 88.0057% and the area under the ROC curve was 0.946. This shows how strong a feature kurtosis really is in determining when birds are stressed by heat or ammonia levels.

CHAPTER VI

CONCLUSIONS

In this thesis the possibility of using audio to monitor bird stress by means of audio recording was analyzed and evaluated. It was found that by extracting the kurtosis and the loudness of said audio, very accurate results could be achieved. It was shown by means of the analysis of a simple threshold classifier over each feature that the kurtosis and the loudness both prooved to be very strong features for detecting stress induced by heat and ammonia. Further, by building an Adaboost classifier to combine all the features into a single classifier, it was shown that using just these two features did not significantly dropped the quality of the classifier as opposed to using all the features.

The classifier model based on these two features proves to be very useful and practical for further tests in the industrial sector. These features can be easily extracted and refined on an embedded system and almost any classifier model (decision trees, boosting) can be utilized to make decisions with regards to stress.

It was observed that birds had one of two reactions when presented with a stressful situation, either they calmed down and spreaded out, as to accept their inevitable fate or they moved and clucked desperately.

Now that a baseline has been established, experiments can be better designed to further improve the capacity of classification of the features found. There are also other types of stress that would be interesting to study.

APPENDIX A

TERMINOLOGY

Most of these terms are defined including information pulled from Wikipedia and other internet sources.

Adrenocorticotropic hormone (ACTH). Also know as ‘corticotropin’, it is a polypeptide tropic hormone produced and secreted by the pituitary gland. It is an important component of the hypothalamic-pituitary-adrenal axis and is often produced in response to biological stress. Its principal effects are increased production and release of corticosteroids and, as its name suggests, cortisol from the adrenal cortex [15].

Mel frequency cepstral coefficients (MFCC). It is a representation of the short-term power spectrum of a sound, based on a linear cosine transform of a log power spectrum on a nonlinear mel scale of frequency [19]. Typically the feature is computed over small windows of time, which are then combined and the first few moments of the combined data are used for classification.

Short-term Fourier transform (STFT). It is a Fourier-related transform used to determine the sinusoidal frequency and phase content of local sections of a signal as it changes over time, over a short window of time [23].

Redundant Array of Independent Disks Level 5 (RAID 5). It is a storage technology that combines multiple disk drive components into a logical unit. Data is distributed across the drives in one of several ways called “RAID levels,” depending on what level of redundancy and performance (via parallel communication) is required. Level 5 (block-level striping with distributed parity) distributes parity along with the data and requires all drives but one to be present to operate; the array is not destroyed by a single drive failure. Upon drive failure, any subsequent reads can be calculated from the distributed parity such that the drive failure is masked from the end user. However, a single drive failure results in reduced performance of the entire array until the failed drive has been replaced and the associated data rebuilt [21].

Revolutions Per Minute (RPM). Term used to describe the speed of a hard drive in getting to certain locations on the magnetic plate.

Universal Serial Bus (USB). It is an industry standard developed in the mid-1990s that defines the cables, connectors and communications protocols used in a bus for connection, communication and power supply between computers and electronic devices [25].

Terabyte (TB). A multiple of the unit byte for digital information, it equals 1,000,000,000,000 bytes.

Advanced Linux Sound Architecture (ALSA). ALSA provides audio and MIDI functionality to the Linux operating system [2].

Comma-Separated Values (CSV). It is a file format that stores tabular data (numbers and text) in plain-text form. Plain text means that the file is a sequence of characters, with no data that has to be interpreted instead, as binary numbers. A CSV file consists of any number of records, separated by line breaks of some kind; each record consists of fields, separated by some other character or string, most commonly a literal comma or tab. Usually, all records have an identical sequence of fields [16].

Receiver operating characteristic (ROC). In signal detection theory, a receiver operating characteristic (ROC), or simply ROC curve, is a graphical plot which illustrates the performance of a binary classifier system as its discrimination threshold is varied. It is created by plotting the fraction of true positives out of the positives (TPR = true positive rate) vs. the fraction of false positives out of the negatives (FPR = false positive rate), at various threshold settings. (TPR is also known as sensitivity, and FPR is one minus the specificity or true negative rate) [22].

United States Department of Agriculture (USDA). The United States Department of Agriculture (informally the Agriculture Department or USDA) is the United States federal executive department responsible for developing and executing U.S. federal government policy on farming, agriculture, and food [24].

Within the USDA, for more than 40 years, Congress has entrusted the Animal and Plant Health Inspection Service (APHIS) with the stewardship of animals covered under the Animal Welfare and Horse Protection Acts. APHIS continues to uphold that trust, giving protection to millions of animals each year, nationwide. APHIS provides leadership for determining standards of humane care and treatment of animals. APHIS implements those standards and achieves compliance through inspection, education, cooperative efforts, and enforcement [13].

Office of Laboratory Animal Welfare (OLAW). The Office of Laboratory Animal Welfare (OLAW) provides guidance and interpretation of the Public Health Service (PHS) Policy on Humane Care and Use of Laboratory Animals, supports educational programs, and monitors compliance with the Policy by Assured institutions and PHS funding components to ensure the humane care and use of animals in PHS-supported research, testing, and training, thereby contributing to the quality of PHS-supported activities [10].

Institutional Animal Care and Use Committee (IACUC). The Institutional

Animal Care and Use Committee (IACUC) is a self-regulating entity that, according to U.S. federal law, must be established by institutions that use laboratory animals for research or instructional purposes to oversee and evaluate all aspects of the institution's animal care and use program [5].

Sound eXchange (SoX). SoX is a cross-platform (Windows, Linux, MacOS X, etc.) command line utility that can convert various formats of computer audio files in to other formats. It can also apply various effects to these sound files, and, as an added bonus, SoX can play and record audio files on most platforms [12].

cron. cron is the time-based job scheduler in Unix-like computer operating systems. cron enables users to schedule jobs (commands or shell scripts) to run periodically at certain times or dates [17].

ncurses. ncurses (new curses) is a programming library that provides an API which allows the programmer to write text-based user interfaces in a terminal-independent manner. It is a toolkit for developing "GUI-like" application software that runs under a terminal emulator. It also optimizes screen changes, in order to reduce the latency experienced when using remote shells [20].

Just-noticeable difference (jnd). In psychophysics, a just noticeable difference, customarily abbreviated with lowercase letters as jnd, is the smallest detectable difference between a starting and secondary level of a particular sensory stimulus. It is also known as the difference limen or the differential threshold [18].

REFERENCES

[1] "How to understand heat stress and what to do about it."

[2] ALSA, "Introduction," 2012. [Online; accessed 8-May-2012].

[3] Beck, J., "Understanding and minimizing stress in broilers and broiler breeders," *Zootecnica International*, vol. XIV(3), pp. 30–38, 1991.

[4] Brake, J., "Stress and flock health," Nov. 1985.

[5] for Laboratory Animal Science, A. A., "About us," 2012. [Online; accessed 8-May-2012].

[6] Freeman, B., "The stress syndrome," *World's Poultry Sci. J.*, vol. 43, pp. 15–19, 1987.

[7] Gross, W., "Effects of stress on poultry," *Poultry Digest*, vol. 45(2), p. 58, 1985.

[8] Gross, W. and Siegel, P., "Long-term exposure of chickens to three levels of social stress," *Avian Dis*, vol. 25, pp. 312–325, 1981.

[9] Miles, D., "Atmospheric ammonia is detrimental to the performance of modern commercial broilers," *Poultry Science*, vol. 83, pp. 1650–1654, July 2004.

[10] of Health, U. D. and Services, H., "Olaw mission statement," 2012. [Online; accessed 8-May-2012].

[11] Rosales, A. G., "Managing stress in broiler breeders: A review," *Applied Poultry Science*, vol. 3, pp. 199–207, 1994.

[12] SoX, "Homepage," 2012. [Online; accessed 8-May-2012].

[13] USDA, "Animal welfare," 2012. [Online; accessed 8-May-2012].

[14] Wichern, G., "Segmentation, indexing, and retrieval for environmental and natural sounds," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, pp. 688–707, Mar. 2010.

[15] Wikipedia, "Adrenocorticotropic hormone," 2012. [Online; accessed 8-May-2012].

[16] Wikipedia, "Comma-separated values," 2012. [Online; accessed 8-May-2012].

[17] Wikipedia, "cron," 2012. [Online; accessed 8-May-2012].

[18] Wikipedia, "Just-noticeable difference," 2012. [Online; accessed 8-May-2012].

[19] Wikipedia, "Mel-frequency cepstrum," 2012. [Online; accessed 8-May-2012].

[20] Wikipedia, "ncurses," 2012. [Online; accessed 8-May-2012].

[21] Wikipedia, "Raid," 2012. [Online; accessed 8-May-2012].

[22] Wikipedia, "Receiver operating characteristic," 2012. [Online; accessed 8-May-2012].

[23] Wikipedia, "Short-time fourier transfrom," 2012. [Online; accessed 8-May-2012].

[24] Wikipedia, "United states department of agriculture," 2012. [Online; accessed 8-May-2012].

[25] Wikipedia, "Universal serial bus," 2012. [Online; accessed 8-May-2012].

Avian Musing Feature Space Analysis

Guillermo J. Colón

48 Pages

Directed by Professor David V Anderson

The purpose of this study was to analyze the possibility of utilizing known signal processing and machine learning algorithms to correlate environmental data to chicken vocalizations. The specific musing to be analyzed consist of not just one chicken's vocalizations but of a whole collective, it therefore becomes a chatter problem. There have been similar attempts to create such a correlation in the past but with singled out birds instead of a multitude. This study was performed on broiler chickens (birds used in meat production).

One of the reasons why this correlation is useful is for the purpose of an automated control system. Utilizing the chickens own vocalization to determine the temperature, the humidity, the levels of ammonia among other environmental factors, reduces, and might even remove, the need for sophisticated sensors.

Another factor that this study wanted to correlate was stress in the chickens to their vocalization. This has great implications in animal welfare, to guarantee that the animals are being properly take care off. Also, it has been shown that the meat of non-stressed chickens is of much better quality than the opposite.

The audio was filtered and certain features were extracted to predict stress. The features considered were loudness, spectral centroid, spectral sparsity, temporal sparsity, transient index, temporal average, temporal standard deviation, temporal skewness, and temporal kurtosis. In the end, out of all the features analyzed it was shown that the kurtosis and loudness proved to be the best features for identifying stressed birds in audio.